Preparations

1. Make sure that the received json file and annotate.py file are within the same folder on your computer.
2. Make sure the required software is installed, if this is done then you should be able to run the python file in order to start the annotation process.

Software Requirements

In order for the annotation script to work the following software has to be installed on your computer:

- Python 3
- Pandas package for python

The easiest way to install both of these is to install the Anaconda package available here https://docs.continuum.io/anaconda/install/. However this does also install several other python packages.

If you instead want to only install Python and the Pandas package download Python from https://www.python.org/, a guide for installing python can be found here https://realpython.com/installing-python/.

In order to install Pandas you first have to install Pip. The following link provides instructions on how to install Pip: https://pip.pypa.io/en/stable/installing/

Pandas can thereafter be installed by running `pip install pandas` in the command prompt of your system.

Goal

The goal of the annotation process will be to link individuals detected in 18th century notary documents to individuals in a large database of VOC employees.

Guidelines

Each entry will show you the following information:

	Notary Information (index number)	VOC Information (index number)
Name	(Notary Name)	(Original Name) / (Normalized Name)
Dates	(Notary Date)	(Leave Date) / (Return Date)
Ships	(Found Ships)	(Leave Ship) / (Return Ship)
Rank	(Found Rank)	(Rank)
Locations	(Found Location)	(Place of Origin)

Everything between brackets will be replaced by the actual information of the individuals that will be compared. When annotating please follow the following steps:

1. Check if the (Notary Name) matches either the (Original Name) or the (Normalized Name). To counteract missing a match due to misspelling of the name, matches with similar names are also shown. Some of these names are actual different names and not simply a case of different spelling.
   An example of a misspelled name would be: Hendrik Volkerts → Hendrik Volkers
   An example of a different name would be: Jan van Dam → Jan van Dalen

2. Compare the (Notary Date) to the (Leave Date) / (Return Date), the closer the (Notary Date) is to either one of the dates, the higher the chance the entries are a match.

3. Compare the (Found Ships) to the (Leave Ship) and (Return Ship), if at least one ship matches this is a strong indicator that the entries are a match. If both ships match, then it is a very strong indicator that the entries are a match.

4. Compare (Found Rank) to (Rank), if they match this is a moderate indicator that the entries are a match.

5. Compare (Found Location) to (Place of Origin), any match aside from “Amsterdam” is a moderate indicator that the entries are a match.

6. If Name, Date, Ship, Rank, and Location match then it can be assumed that the entries are a match, type y and hit enter to confirm this. If this is not the case go to step 7.

7. If Name and Date match, type text and hit enter. This will show you the Handwritten Text Recognition text of the notary entry. While this text might be very hard to understand, please scan the text for variations of the following keywords which might indicate involvement with the VOC: “Schip”, “Oostindische Compagnie”, “Oostindie”, “Kamer Amsterdam”, “Kamer Zeeland”.
   If there are any matches in Ship, Rank, or Location, then please also look for these words.

8. If after scanning the text you think the entries are match, type y and hit enter. If not type n and hit enter.

9. Once all entries have been annotated the program will shut down. Please send the final.json file that is now present in the same folder as the annotate.py file back.

Possible Issues

What to do when you are uncertain whether an entry is a match or not:
Please note down the name and index number of both the Notary information and VOC information. Assign this entry as a match and once you are done annotating provide the recorded name and index numbers along with the final.json file.

What to do if you want to stop annotating:
If you would like to stop annotating please enter stop and hit enter. This will save your current progress in the final.json file. If you would later like to continue annotating, you can simply run the annotate.py file again. The script will then check the final.json file for the last entry and allow you to continue annotating from there.

Tips

After scanning the Handwritten Text Recognition text and you have determined the person did not sail for the VOC, please pay attention to the Notary Information index number. You might receive more entries for the same person with the same Notary Information index number. This allows you to immediately confirm that the entries are not a match.